

Quantum information causality

Damián Pitalúa-García

*Centre for Quantum Information and Foundations, DAMTP, Centre for Mathematical Sciences,
University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, U.K.*

How much information can a transmitted physical system fundamentally communicate? We introduce the principle of *quantum information causality*, which states the maximum amount of quantum information that a quantum system can communicate as a function of its dimension, independently of any previously shared quantum physical resources. We present a new quantum information task, whose success probability is upper bounded by the new principle, and show that an optimal strategy to perform it combines the quantum teleportation and superdense coding protocols with a task that has classical inputs.

Quantum information science studies how information can fundamentally be encoded, processed and communicated via systems described by quantum physics [1]. Interesting features of information arise with this approach. The no-cloning theorem states that unknown quantum states cannot be copied perfectly [2, 3]. Unknown quantum states can be teleported [4]. Two classical bits can be encoded in one qubit via the superdense coding protocol [5]. Fundamentally-secure cryptography can be achieved with quantum information protocols [6–8]. Many of the quantum information protocols are possible due to quantum entanglement: two systems are entangled if their global quantum state cannot be expressed as a convex combination of individual states in a tensor product form. Another interesting property is quantum nonlocality, that is, measurement outcomes of separate systems can exhibit correlations that cannot be described by local classical models [9, 10].

Since the value of quantum correlations does not vary with the time difference of the measurements and the distance between the systems, one could think that they can be used to communicate arbitrarily-fast messages. However, quantum physics obeys the no-signaling principle. No-signaling says that a measurement outcome obtained by a party (Bob) does not provide him with any information about what measurement is performed by another party (Alice) at a distant location, despite any nonlocal correlations previously shared by them [11].

If any information that Alice has is to be learned by Bob, no-signaling requires that a physical system sharing correlations with Alice’s system must be transmitted to him. Thus, an interesting question to ask is: how much information can a physical system fundamentally communicate? In the scenario in which Alice has a classical random variable X , she encodes its value in a quantum state that she sends Bob and Bob applies a quantum measurement on the received state in order to obtain a classical random variable Y as the output, the Holevo theorem [12] provides an upper bound on the classical mutual information between X and Y . In the scenario in which Alice sends Bob m classical bits, information causality states that the increase of the classical mutual

information between Bob’s and Alice’s systems is upper bounded by m , independently of any no-signaling physical resources that Alice and Bob previously shared [13]. Information causality has important implications for the set of quantum correlations [13–17]. For example, it implies the Cirel’son bound [18], while the no-signaling principle does not [19].

Here we consider the scenario in which Bob receives a quantum system from Alice, who possibly shares quantum correlations with another party, Charlie, and ask the question: how much quantum information can Bob obtain about Alice’s or Charlie’s data? We introduce a new principle that we call *quantum information causality*, which states that the maximum amount of quantum information that a quantum system can communicate is limited by its dimension, independently of any quantum physical resources previously shared by the communicating parties. Namely, the principle says that *the increase of the quantum mutual information between Bob’s and Charlie’s systems, after a quantum system of m qubits is transmitted from Alice to Bob, is upper bounded by $2m$.*

In order to illustrate quantum information causality, we introduce a new quantum task that we call the *quantum information causality (QIC) game*. For convenience we present two equivalent versions of it.

The QIC game (version I). Initially, Alice and Bob may share an arbitrary entangled state. However, they do not share any correlations with Charlie. Let B denote the quantum system at Bob’s site. Charlie prepares n qubits in the pure states $\{|\psi_j\rangle\}_{j=0}^{n-1}$, completely randomly. Charlie sends Alice the qubit A_j in the quantum state $|\psi_j\rangle$, for $j = 0, 1, \dots, n-1$, and keeps a classical record of the states. We denote the global system that Alice receives from Charlie as $A \equiv A_0 A_1 \dots A_{n-1}$. Charlie generates a random integer $k \in \{0, 1, \dots, n-1\}$ and gives it to Bob. Bob gives Charlie a qubit B_k in the state ρ_k , which must be as close as possible to $|\psi_k\rangle$. Alice and Bob may play any strategy allowed by quantum physics as long as the following constraint is satisfied: their communication is limited to a single message from Alice to Bob only, encoded in a quantum system T of $m < n$ qubits, with no extra classical communication allowed.

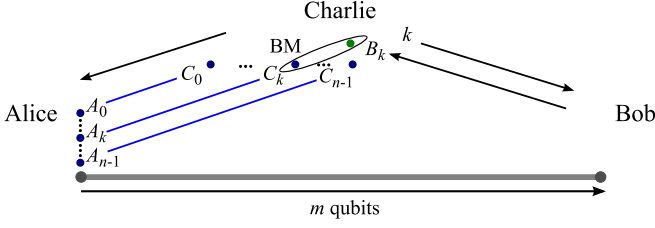


FIG. 1. The QIC game (version II).

Charlie measures the received state ρ_k in the orthonormal basis $\{|\psi_k\rangle, |\psi_k^\perp\rangle\}$, where $|\psi_k^\perp\rangle$ is the qubit state with Bloch vector antiparallel to that one of $|\psi_k\rangle$. Alice and Bob win the game if Charlie's measurement outcome corresponds to the state $|\psi_k\rangle$. The success probability is

$$p \equiv \int d\mu_0 \int d\mu_1 \cdots \int d\mu_{n-1} \left(\frac{1}{n} \sum_{k=0}^{n-1} \langle \psi_k | \rho_k | \psi_k \rangle \right), \quad (1)$$

where $\int d\mu_j$ is the normalized integral over the Bloch sphere corresponding to the state $|\psi_j\rangle$.

The QIC game (version II). This version is similar to version I with the following differences (see Fig. 1). Charlie does not prepare pure qubit states. Instead, Charlie prepares the qubits A_j and C_j in the singlet state $|\Psi^-\rangle$, for $j = 0, 1, \dots, n-1$. Charlie keeps the system $C \equiv C_0 C_1 \cdots C_{n-1}$ and sends Alice the system $A \equiv A_0 A_1 \cdots A_{n-1}$. We denote the joint state of the qubits C_k and B_k as ω_k . Charlie applies a Bell measurement (BM) on the joint system $C_k B_k$. Alice and Bob win the game if Charlie obtains the outcome corresponding to the singlet state. The success probability is

$$P \equiv \frac{1}{n} \sum_{k=0}^{n-1} \langle \Psi^- | \omega_k | \Psi^- \rangle. \quad (2)$$

We show in the Supplemental Material that both versions of the QIC game are equivalent and that their success probabilities satisfy the relation $p = (1 + 2P)/3$. It is more convenient to work with version II of the QIC game. Thus, we only refer to this version in the rest of this Letter, unless otherwise stated.

Consider the following *naive* strategy to play the QIC game. Alice simply sends Bob m of the n received qubits from Charlie without applying any operation on these. Alice and Bob previously agree on which qubits Alice would send Bob, for example, those with index $0 \leq j < m$. If Bob receives from Charlie a number $k < m$, he outputs the correct state; in this case, $\langle \Psi^- | \omega_k | \Psi^- \rangle = 1$. However, if $m \leq k$, Bob does not have the correct state, hence, he can only give Charlie a fixed state, say $|0\rangle$; in this case, $\langle \Psi^- | \omega_k | \Psi^- \rangle = 1/4$. Thus, this strategy succeeds with probability

$$P_N = \frac{1}{4} \left(1 + 3 \frac{m}{n} \right), \quad (3)$$

where the label N stands for *naive*. There are other strategies that achieve success probabilities higher than P_N . However, it turns out that in general, $P < 1$, if $m < n$. We show that this is a consequence of quantum information causality.

The principle of *quantum information causality* states an upper bound on the amount of quantum information that m qubits can communicate:

$$\Delta I(C : B) \leq 2m, \quad (4)$$

where $\Delta I(C : B) \equiv I(C : B') - I(C : B)$ is Bob's gain of quantum information about C , $I(C : B) \equiv S(C) + S(B) - S(CB)$ is the quantum mutual information [1] between C and B , $S(C)$ is the von Neumann entropy [1] of C , etc., B' denotes the joint system BT after Bob's quantum operations.

The proof is very simple. By definition, $I(C : BT) = S(C) + S(BT) - S(CBT)$. Subadditivity [20] states that $S(BT) \leq S(B) + S(T)$. The triangle inequality [21], $|S(CB) - S(T)| \leq S(CBT)$, implies that $-S(CBT) \leq S(T) - S(CB)$. Hence, we have that $I(C : BT) \leq 2S(T) + I(C : B)$. The data-processing inequality states that local operations can only decrease the quantum mutual information [1]. Thus, $I(C : B') \leq I(C : BT)$, which implies that $I(C : B') \leq 2S(T) + I(C : B)$. Therefore, we obtain that $\Delta I(C : B) \leq 2S(T)$. Finally, since $S(T) \leq \log_2(\dim T)$, the quantum information that T can communicate is limited by its dimension. Therefore, if T is a system of m qubits, Eq. (4) follows because in this case $S(T) \leq m$. Achievability of equality in Eq. (4) requires that T is maximally entangled with C (see details in the Supplemental Material). It is easy to see that the naive strategy in the QIC game saturates this bound.

Information causality is obtained as a special case of quantum information causality. If the transmitted system T is classical, equality in Eq. (4) can no longer be achieved. Information causality states that in this case, $\Delta I(C : B) \leq m$, where C and B are classical systems and $I(C : B)$ denotes their classical mutual information. In fact, this bound is valid even if C and B are quantum and $I(C : B)$ denotes the quantum mutual information (see details in the Supplemental Material).

As stated above, quantum information causality follows from three properties of the von Neumann entropy: subadditivity, the data-processing and the triangle inequalities. The concept of entropy in mathematical frameworks for general probabilistic theories [22–24] and its implication for information causality have been recently investigated [25–27]. Particularly, it has been shown that a physical condition on the measure of entropy implies subadditivity and the data-processing inequality, and hence that information causality follows from this condition [27]. It would be interesting to investigate whether physically-sensible definitions of entropy for more general probabilistic theories satisfy the three mentioned properties, and hence a generalized version of

quantum information causality. A different version of information causality in more general probabilistic theories has been considered in Ref. [28].

Quantum information causality implies an upper bound on the success probability in the QIC game:

$$P \leq P', \quad (5)$$

where we define P' to be the maximum solution of the equation

$$h(P') + (1 - P') \log_2 3 = 2 \left(1 - \frac{m}{n}\right), \quad (6)$$

and $h(x) = -x \log_2 x - (1 - x) \log_2 (1 - x)$ denotes the binary entropy. The value of P' increases monotonically with the ratio m/n , achieving $P' = 1/4$ if $m = 0$ and $P' = 1$ if $m = n$. Therefore, we have that $P < 1$ if $m < n$. Some values of P' are plotted in Fig. 2. The complete proof of Eq. (5) is given in the Supplemental Material. Below we present a sketch of it.

Firstly, we notice that for any strategy that Alice and Bob may play that achieves success probability P , there exists a covariant strategy achieving the same value of P that Alice and Bob can perform. By covariance, we mean the following: if, when Alice's input qubit A_k is in the state $|\psi_k\rangle$, Bob's output qubit state is ρ_k , then, when A_k is in the state $U|\psi_k\rangle$, Bob's output state is $U\rho_k U^\dagger$, for any qubit state $|\psi_k\rangle \in \mathbb{C}^2$ and unitary operation $U \in \text{SU}(2)$. Recall that k is the number that Charlie gives Bob. Therefore, without loss of generality, we consider that a covariant strategy is implemented. In version I of the QIC game, this means that the Bloch sphere of the qubit A_k is contracted uniformly and output in the qubit B_k . In version II, this means that the joint system $C_k B_k$ is transformed into the state

$$\omega_k = \lambda_k \Psi^- + \frac{1 - \lambda_k}{3} (\Psi^+ + \Phi^+ + \Phi^-), \quad (7)$$

where $1/4 \leq \lambda_k \leq 1$ and Ψ^- denotes $|\Psi^-\rangle\langle\Psi^-|$, etc. That is, the depolarizing map [1] is applied to the qubit A_k , and output by Bob in the qubit B_k .

Then, we use the data-processing inequality and the fact that the qubits C_j and $C_{j'}$ are in a product state for every $j \neq j'$ in order to show that $\sum_{k=0}^{n-1} I(C_k : B_k) \leq I(C : B')$. We notice that since Charlie's and Bob's systems are initially uncorrelated, Eq. (4) reduces to $I(C : B') \leq 2m$. Thus, we have that $\sum_{k=0}^{n-1} I(C_k : B_k) \leq 2m$. From this inequality and the concavity property of the von Neumann entropy, we obtain an upper bound on $\sum_{k=0}^{n-1} \lambda_k/n$, which from Eqs. (2) and (7) equals P .

Below we show that an optimal strategy to play the QIC game reduces to an optimal strategy to perform the following task.

The IC-2 game. Alice is given random numbers $x_j \equiv (x_j^0, x_j^1)$, where $x_j^0, x_j^1 \in \{0, 1\}$, for $j = 0, 1, \dots, n-1$. Bob is given a random value of $k = 0, 1, \dots, n-1$. The game's

goal is that Bob outputs x_k . Alice and Bob can perform any strategy allowed by quantum physics with the only condition that communication is limited to a single message of $2m < 2n$ bits from Alice to Bob. In particular, Alice and Bob may share an arbitrary entangled state. Let $y_k \equiv (y_k^0, y_k^1)$ be Bob's output, where $y_k^0, y_k^1 \in \{0, 1\}$. We define the success probability as

$$Q \equiv \frac{1}{n} \sum_{k=0}^{n-1} P(y_k = x_k). \quad (8)$$

We call this task the *IC-2* game. The version we call the *IC-1* game, in which the inputs and output are one bit values and Alice's message is of $m < n$ bits, was considered in the paper that introduced information causality [13]. The strategies to play the IC-1 game in which no entanglement is used were first considered by Wiesner in 1983 with the name of conjugate coding [29]. They were investigated further in 2002 with the name of random access codes (RACs) [30]. The most general quantum strategy, in which Alice and Bob share an arbitrary entangled state, is called an entanglement-assisted random access code (EARAC) [31].

Let Q_{\max} be the maximum value of Q over all possible strategies to play the IC-2 game. We show that

$$P \leq Q_{\max}. \quad (9)$$

Consider the following strategy to play the IC-2 game. Alice and Bob initially share a singlet state in the qubits A_j and C_j , for $j = 0, 1, \dots, n-1$. Alice has the system $A \equiv A_0 A_1 \dots A_{n-1}$, while Bob has the system $C \equiv C_0 C_1 \dots C_{n-1}$. Alice applies the unitary operation σ_{x_j} on the qubit A_j , for every j , where $\sigma_{0,0} \equiv I$ is the identity operator acting on \mathbb{C}^2 and $\sigma_{0,1} \equiv \sigma_1$, $\sigma_{1,0} \equiv \sigma_2$, $\sigma_{1,1} \equiv \sigma_3$ are the Pauli matrices. Then, Alice and Bob play the QIC game, applying some operation on the input system A , which includes a message of m qubits from Alice to Bob. However, instead of sending these m qubits directly, Alice teleports [4] them to Bob. Thus, communication consists of $2m$ bits only, as required. At this stage, Bob does not apply any operation on the system C , which is consistent with the QIC game. As it was previously indicated, we can consider that in a general strategy in the QIC game the depolarizing map is applied to the qubit A_k (see Eq. (7)). Therefore, Bob outputs the qubit B_k in the following joint state with the qubit C_k :

$$\Omega_k = I \otimes \sigma_{x_k} \left[\lambda_k \Psi^- + \frac{1 - \lambda_k}{3} (\Psi^+ + \Phi^+ + \Phi^-) \right] I \otimes \sigma_{x_k}.$$

Then, Bob measures the state Ω_k in the Bell basis. Bob learns the encoded value x_k with probability λ_k . Thus, from Eq. (8) we have that $Q = \sum_{k=0}^{n-1} \lambda_k/n$, which equals P , as we can see from Eqs. (2) and (7). Since by definition $Q \leq Q_{\max}$, we have that $P \leq Q_{\max}$, as claimed.

Consider the following class of strategies to play the QIC game that combine quantum teleportation [4], superdense coding [5] (SDC) and the IC-2 game.

Teleportation strategies in the QIC game. Alice and Bob share a singlet state in the qubits A'_j , at Alice's site, and B_j , at Bob's site, for $j = 0, 1, \dots, n-1$. Alice applies a Bell measurement on her qubits $A_j A'_j$ and obtains the two bit outcome $x_j \equiv (x_j^0, x_j^1)$. Thus, the state of the qubit A_j is teleported to Bob's qubit B_j , up to the Pauli error σ_{x_j} . This means that the joint state of the system $C_j B_j$ transforms into one of the four Bell states, according to the value of x_j . Alice and Bob play the IC-2 game with Alice's and Bob's inputs being $x \equiv (x_0, x_1, \dots, x_{n-1})$ and k , respectively. However, instead of sending Bob the $2m$ -bits message directly, Alice encodes it in m qubits via SDC. Bob receives the m qubits and decodes the correct $2m$ -bits message, which he inputs to his part of the IC-2 game. Bob outputs the two bit number $y_k \equiv (y_k^0, y_k^1)$ and applies the Pauli correction operation σ_{y_k} on the qubit B_k , which then he outputs and gives to Charlie. If $y_k = x_k$, the output state ω_k of the joint system $C_k B_k$ is the singlet; otherwise, we have that $\langle \Psi^- | \omega_k | \Psi^- \rangle = 0$. Thus, from the definition of P , Eq. (2), we see that $P = Q$, where Q is given by Eq. (8).

Therefore, from Eq. (9) we see that an optimal strategy in the QIC game is a teleportation strategy, as defined above, in which the IC-2 game is played achieving the maximum success probability $Q = Q_{\max}$. We have obtained an upper bound on Q for a particular class of strategies in the case $m = 1$ (see Supplemental Material).

The best strategy that we have found to play the QIC game in the case $m = 1$ is a teleportation strategy in which the IC-2 game is played with two equivalent and independent protocols in the IC-1 game. In both protocols Bob inputs the number k , while Alice inputs the bits $\{x_j^0\}_{j=0}^{n-1}$ in the first protocol and the bits $\{x_j^1\}_{j=0}^{n-1}$ in the second one. If Bob outputs the correct value of x_k^0 with probability q in the first protocol, and similarly, he outputs the correct value of x_k^1 with probability q in the second protocol, for any k , then the success probability in the IC-2 game is $Q = q^2$. The maximum value of q that has been shown [27, 31] is $q = (1 + n^{-1/2})/2$. Explicit strategies to achieve this value are given by EARACs in the case in which $n = 2^r 3^l$ and r, l are nonnegative integers [31]. With this value of Q we achieve a success probability in the QIC game of

$$P_T = \frac{1}{4} \left(1 + \frac{1}{\sqrt{n}} \right)^2, \quad (10)$$

where the label T stands for *teleportation*.

Here we have introduced the quantum information causality principle as satisfaction of an upper bound on the quantum information that Bob can gain about Charlie's data as a function of the number of qubits m that Alice (who shares correlations with Charlie) sends Bob, Eq. (4). We have presented a new quantum information task, the QIC game, whose success probability is limited by quantum information causality, Eq. (5). We have

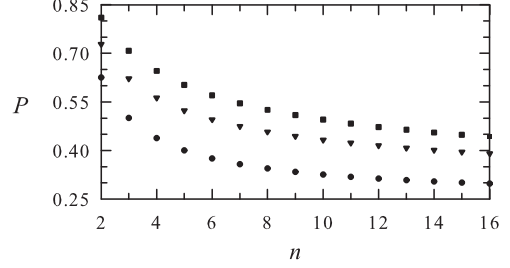


FIG. 2. Success probability (P) in the QIC game for $m = 1$ achieved with the naive strategy, P_N (circles), and with the best teleportation strategy that we have found, P_T (triangles). The upper bound on P obtained from quantum information causality, P' (squares), is plotted too.

shown that an optimal strategy to play the QIC game combines the quantum teleportation and the quantum superdense coding protocols, with an optimal strategy to perform another task that has classical inputs, the IC-2 game. An optimal strategy in the IC-2 game remains as an interesting open problem.

I would like to thank Adrian Kent for much assistance with this work, and Nilanjana Datta, Sabri Al-Safi, Tony Short and Min-Hsiu Hsieh for helpful discussions. I acknowledge financial support from CONACYT México and partial support from Gobierno de Veracruz.

SUPPLEMENTAL MATERIAL

Equivalence between the two versions of the QIC game

We show that both versions of the QIC game are equivalent and that their success probabilities satisfy the relation $p = (1 + 2P)/3$. More precisely, we show that if Alice and Bob play a strategy in version I of the QIC game that achieves a success probability p , the same strategy applied to version II achieves a success probability P that satisfies the relation $p = (1 + 2P)/3$, for any strategy that they may play, and vice versa.

We change to a more convenient notation, $|\psi_k\rangle \equiv |\uparrow_{\vec{r}_k}\rangle$, $|\psi_k^\perp\rangle \equiv |\downarrow_{\vec{r}_k}\rangle$, in order to make clear that $|\psi_k\rangle$ and $|\psi_k^\perp\rangle$ correspond to pure qubit states with Bloch vectors \vec{r}_k and $-\vec{r}_k$, respectively.

Version I of the QIC game is equivalent to the following. Charlie initially prepares the pair of qubits A_j and C_j in the singlet state $|\Psi^-\rangle$, he gives Alice the qubit A_j and keeps the qubit C_j , for $j = 0, 1, \dots, n-1$. Charlie generates a random integer $k \in \{0, 1, \dots, n-1\}$ and gives it to Bob. Charlie measures the joint state ω_k of his qubit C_k and the one received by Bob B_k in the orthonormal

basis $\mathcal{B}_{\vec{r}_k} \equiv \{|\uparrow_{\vec{r}_k}\rangle|\uparrow_{\vec{r}_k}\rangle, |\downarrow_{\vec{r}_k}\rangle|\downarrow_{\vec{r}_k}\rangle, |\uparrow_{\vec{r}_k}\rangle|\downarrow_{\vec{r}_k}\rangle, |\downarrow_{\vec{r}_k}\rangle|\uparrow_{\vec{r}_k}\rangle\}$ for some vector \vec{r}_k that he chooses completely randomly from the Bloch sphere. Opposite outcomes correspond to success. Therefore, the success probability p that Alice and Bob achieve in version I of the QIC game, given by Eq. (1) of the main text, equals the following in this version:

$$p = \int d\mu_0 \int d\mu_1 \cdots \int d\mu_{n-1} \left[\frac{1}{n} \sum_{k=0}^{n-1} (\langle \uparrow_{\vec{r}_k} | \langle \downarrow_{\vec{r}_k} | \omega_k | \uparrow_{\vec{r}_k} \rangle | \downarrow_{\vec{r}_k} \rangle + \langle \downarrow_{\vec{r}_k} | \langle \uparrow_{\vec{r}_k} | \omega_k | \downarrow_{\vec{r}_k} \rangle | \uparrow_{\vec{r}_k} \rangle) \right], \quad (11)$$

where $\int d\mu_j$ is the normalized integral over the Bloch sphere corresponding to the Bloch vector \vec{r}_j .

The Bell states defined in the basis $\mathcal{B}_{\vec{r}_k}$ are

$$\begin{aligned} |\Phi_{\vec{r}_k}^{\pm}\rangle &\equiv \frac{1}{\sqrt{2}} (|\uparrow_{\vec{r}_k}\rangle|\uparrow_{\vec{r}_k}\rangle \pm |\downarrow_{\vec{r}_k}\rangle|\downarrow_{\vec{r}_k}\rangle), \\ |\Psi_{\vec{r}_k}^{\pm}\rangle &\equiv \frac{1}{\sqrt{2}} (|\uparrow_{\vec{r}_k}\rangle|\downarrow_{\vec{r}_k}\rangle \pm |\downarrow_{\vec{r}_k}\rangle|\uparrow_{\vec{r}_k}\rangle). \end{aligned}$$

Consider that instead of measuring the state ω_k in the basis $\mathcal{B}_{\vec{r}_k}$, Charlie measures it in this Bell basis. Since the singlet state is the same in any basis, this corresponds to version II of the QIC game. Therefore, versions I and II of the QIC game are equivalent. Below we show that their success probabilities satisfy the claimed relation.

Using the Bell basis, we obtain from Eq. (11) that

$$p = \int d\mu_0 \int d\mu_1 \cdots \int d\mu_{n-1} \left[\frac{1}{n} \sum_{k=0}^{n-1} (\langle \Psi_{\vec{r}_k}^{-} | \omega_k | \Psi_{\vec{r}_k}^{-} \rangle + \langle \Psi_{\vec{r}_k}^{+} | \omega_k | \Psi_{\vec{r}_k}^{+} \rangle) \right]. \quad (12)$$

Since the singlet state $|\Psi_{\vec{r}_k}^{-}\rangle$ is the same in any basis, by the definition of P (Eq. (2) of the main text), we have that

$$\int d\mu_0 \int d\mu_1 \cdots \int d\mu_{n-1} \frac{1}{n} \sum_{k=0}^{n-1} \langle \Psi_{\vec{r}_k}^{-} | \omega_k | \Psi_{\vec{r}_k}^{-} \rangle = P. \quad (13)$$

On the other hand, we have that

$$\begin{aligned} &\int d\mu_0 \int d\mu_1 \cdots \int d\mu_{n-1} \langle \Psi_{\vec{r}_k}^{+} | \omega_k | \Psi_{\vec{r}_k}^{+} \rangle \\ &= \int d\mu_0 \int d\mu_1 \cdots \int d\mu_{n-1} \text{Tr}(\omega_k |\Psi_{\vec{r}_k}^{+}\rangle \langle \Psi_{\vec{r}_k}^{+}|) \\ &= \text{Tr} \left(\int d\mu_0 \int d\mu_1 \cdots \int d\mu_{n-1} \omega_k |\Psi_{\vec{r}_k}^{+}\rangle \langle \Psi_{\vec{r}_k}^{+}| \right) \\ &= \text{Tr} \left(\omega_k \int d\mu_0 \int d\mu_1 \cdots \int d\mu_{n-1} |\Psi_{\vec{r}_k}^{+}\rangle \langle \Psi_{\vec{r}_k}^{+}| \right) \\ &= \text{Tr} \left(\omega_k \int d\mu_k |\Psi_{\vec{r}_k}^{+}\rangle \langle \Psi_{\vec{r}_k}^{+}| \right), \end{aligned} \quad (14)$$

where in the third line we have used the linearity of the trace; in the fourth line we have used the fact that ω_k

does not depend on the Bloch vector \vec{r}_k because Charlie chooses it completely randomly to define the measurement basis $\mathcal{B}_{\vec{r}_k}$, and can do so after Bob gives him the qubit B_k , and naturally does not depend on the Bloch vectors \vec{r}_j with $j \neq k$ for the same reason; and in the last line we have used that the state $|\Psi_{\vec{r}_k}^{\pm}\rangle$ is defined in terms of the Bloch vector \vec{r}_k , which is parameterized by μ_k , and so is independent of the parameters μ_j with $j \neq k$.

It is easy to obtain that

$$\int d\mu_k |\Psi_{\vec{r}_k}^{\pm}\rangle \langle \Psi_{\vec{r}_k}^{\pm}| = \frac{1}{3} (I - |\Psi^{-}\rangle \langle \Psi^{-}|), \quad (15)$$

where $|\Psi^{-}\rangle \equiv (|01\rangle - |10\rangle)/\sqrt{2}$ is the singlet state in the computational basis and I is the identity operator acting on \mathbb{C}^4 . From Eqs. (14), (15) and the definition of P we have that

$$\frac{1}{n} \sum_{k=0}^{n-1} \int d\mu_0 \int d\mu_1 \cdots \int d\mu_{n-1} \langle \Psi_{\vec{r}_k}^{+} | \omega_k | \Psi_{\vec{r}_k}^{+} \rangle = \frac{1}{3} - \frac{1}{3} P. \quad (16)$$

Finally, we substitute Eqs. (13) and (16) into Eq. (12) to obtain that $p = (1 + 2P)/3$, as claimed.

Achievability of the quantum information causality bound

We show that equality in Eq. (4) of the main text, $\Delta I(C : B) \leq 2m$, requires that the transmitted system T is maximally entangled with Charlie's system C .

Following the proof of Eq. (4) of the main text, we note that equality requires the following conditions to be satisfied. The transmitted system T cannot be entangled with Bob's system B in order to satisfy $S(BT) = S(B) + S(T)$. The system T can only be entangled with the joint system CB so that we have $-S(CBT) = S(T) - S(CB)$, as it is shown below. The state of the system T has to be completely mixed so that its entropy is maximum: $S(T) = m$. This means that T has to be maximally entangled with the system that purifies it. Together, these conditions imply that T has to be maximally entangled with C . We also require that the quantum mutual information between BT and C does not decrease by Bob's operations: $I(C : B') = I(C : BT)$.

Now we show that satisfaction of the equation $-S(CBT) = S(T) - S(CB)$ is achieved if and only if T is entangled only with the joint system CB [1]. Let A be the quantum system that Charlie gives Alice, and hence is initially maximally entangled with C . Let any other physical system that Alice has to be denoted by A' . In particular, A' can be entangled with Bob's system B , but not with Charlie's system C . Let T be the system that Alice sends Bob. Since the systems A' and B are arbitrarily big, without loss of generality, we can consider that the global system $AA'CBT$ is in a pure state. Alice applies some quantum operation on the system TAA' ,

which in general can be represented by a unitary operation followed by a projective measurement. Thus, after Alice's operation, the global system $AA'CBT$ remains in a pure state. Due to the Schmidt decomposition of a bipartite pure state, we have that

$$\begin{aligned} S(CB) &= S(TAA'), \\ S(AA') &= S(CBT). \end{aligned} \quad (17)$$

We apply the subadditivity property to obtain

$$S(TAA') \leq S(AA') + S(T), \quad (18)$$

which from Eq. (17) implies that

$$S(CB) \leq S(CBT) + S(T). \quad (19)$$

Equality in Eq. (19) is achieved if and only if equality in Eq. (18) is satisfied, which occurs if and only if T is in a product state with AA' . Therefore, the relation $-S(CBT) = S(T) - S(CB)$ is satisfied if and only if T is entangled only with the system CB , as claimed.

The information causality bound

If the transmitted system T is classical, equality in Eq. (4) of the main text, $\Delta I(C : B) \leq 2m$, can no longer be achieved. If T represents a classical variable of m bits then the smaller upper bound $\Delta I(C : B) \leq m$ is satisfied. The only difference in the proof of this bound compared to the one of $\Delta I(C : B) \leq 2m$ is that if T is classical then the bound $-S(CBT) \leq S(T) - S(CB)$ can no longer be saturated. In fact, in this case the smaller upper bound $-S(CBT) \leq -S(CB)$ is satisfied. A way to see this is that, if T is a classical variable, the state of the joint system CBT is a distribution over all possible values x of T and states of CB for each x . Therefore, there exists a transformation $x \rightarrow (CB)_x$. From the data-processing inequality we have that $I(CB : T) \leq I(T : T)$. Hence, since $I(CB : T) = S(CB) + S(T) - S(CBT)$ and $I(T : T) = S(T)$, we obtain $S(CB) \leq S(CBT)$ [13].

Reduction of a general strategy in the QIC game to a covariant strategy

For convenience, consider version I of the QIC game in which Charlie gives Alice n pure qubits in the product state $\vec{\psi} \equiv \otimes_{j=0}^{n-1} (|\psi_j\rangle\langle\psi_j|)_{A_j} \in \mathcal{D}((\mathbb{C}^2)^{\otimes n})$, where we define $\mathcal{D}(\mathcal{H})$ to be the set of density operators acting on the Hilbert space \mathcal{H} . Let $\Gamma_k : \mathcal{D}((\mathbb{C}^2)^{\otimes n}) \rightarrow \mathcal{D}(\mathbb{C}^2)$ be the map that Alice and Bob apply to the state $\vec{\psi}$, which outputs the state $\rho_k \equiv \Gamma_k(\vec{\psi})$ that Bob gives Charlie. Recall that k is the number that Charlie gives Bob. After averaging over all possible input pure product states of

qubits with index $j \neq k$, the output only depends on the state $\psi_k \equiv |\psi_k\rangle\langle\psi_k|$, which we identify with the map

$$\begin{aligned} \bar{\Gamma}_k(\psi_k) &\equiv \int d\mu_0 \int d\mu_1 \cdots \int d\mu_{k-1} \int d\mu_{k+1} \int d\mu_{k+2} \cdots \int d\mu_{n-1} \Gamma_k(\vec{\psi}), \end{aligned} \quad (20)$$

where $\int d\mu_j$ is the normalized integral over the Bloch sphere corresponding to the state $|\psi_j\rangle$.

We define the map

$$\bar{\Gamma}_k^{\text{cov}}(\phi) \equiv \int d\nu U_\nu^\dagger \bar{\Gamma}_k(U_\nu \phi U_\nu^\dagger) U_\nu, \quad (21)$$

where $\phi \in \mathcal{D}(\mathbb{C}^2)$, $U_\nu \in \text{SU}(2)$ and $d\nu$ is the Haar measure on $\text{SU}(2)$. It is easy to see that this map is covariant, that is, $\bar{\Gamma}_k^{\text{cov}}(U \phi U^\dagger) = U \bar{\Gamma}_k^{\text{cov}}(\phi) U^\dagger$, for all $\phi \in \mathcal{D}(\mathbb{C}^2)$ and $U \in \text{SU}(2)$.

In principle, for any map Γ_k that Alice and Bob perform, they can implement the covariant map $\bar{\Gamma}_k^{\text{cov}}$ as follows. Alice and Bob initially share randomness. With uniform probability, they obtain the random number ν in the range $d\nu$ that corresponds to an, ideally, infinitesimal region of the Haar measure on $\text{SU}(2)$. This can be done, for example, if Alice and Bob share a maximally entangled state of arbitrarily big dimension and they both apply a local projective measurement in the Schmidt basis on their part of the state; their measurement outcome indicates the number ν . Alice applies the unitary operation U_ν parameterized by the obtained number ν on each of her input qubit states $|\psi_j\rangle$. Then, Alice and Bob apply the map Γ_k to the input state $\otimes_{j=0}^{n-1} (U_\nu |\psi_j\rangle\langle\psi_j| U_\nu^\dagger)_{A_j}$. Finally, Bob applies the unitary U_ν^\dagger to his output qubit. From Eq. (20) we obtain that, after averaging over all possible input pure qubits states with index distinct to k and after Bob's final unitary operation U_ν^\dagger , Bob's output state is $U_\nu^\dagger \bar{\Gamma}_k(U_\nu \psi_k U_\nu^\dagger) U_\nu$. Averaging over all shared random numbers ν , we obtain $\bar{\Gamma}_k^{\text{cov}}(\psi_k)$, as defined by Eq. (21).

It is straightforward to see that the map $\bar{\Gamma}_k^{\text{cov}}$ satisfies

$$\int d\mu_k \langle \psi_k | \bar{\Gamma}_k^{\text{cov}}(\psi_k) | \psi_k \rangle = \int d\mu_k \langle \psi_k | \bar{\Gamma}_k(\psi_k) | \psi_k \rangle.$$

Therefore, it achieves the same value of p (see Eq. (1) of the main text) as $\bar{\Gamma}_k$. Thus, by convenience we consider that Alice and Bob implement the covariant map $\bar{\Gamma}_k^{\text{cov}}(\psi_k)$. In general, this is the depolarizing map [1]:

$$\bar{\Gamma}_k^{\text{cov}}(\phi) = \sum_{i=0}^3 E_i \phi E_i^\dagger,$$

where $\phi \in \mathcal{D}(\mathbb{C}^2)$, $E_0 = \lambda_k I$, $E_i = ((1 - \lambda_k)/3) \sigma_i$, $1/4 \leq \lambda_k \leq 1$ and σ_i are the Pauli matrices, for $i = 1, 2, 3$. Application of the depolarizing map to a qubit that is in the singlet state with another qubit, as in version II of the QIC game, gives as output the state ω_k given by Eq. (7) of the main text.

A useful bound

We show the bound

$$\sum_{k=0}^{n-1} I(C_k : B_k) \leq I(C : B'), \quad (22)$$

which will be useful to deduce an upper bound on P . The proof is equivalent to the one for classical bits [13].

We notice that

$$\begin{aligned} I(C : B') &\equiv I(C_0 C_1 \cdots C_{n-1} : B') \\ &= I(C_0 : B') + I(C_1 C_2 \cdots C_{n-1} : B' C_0) \\ &\quad - I(C_1 C_2 \cdots C_{n-1} : C_0). \end{aligned} \quad (23)$$

Since Charlie's qubits are in a product state with each other, we have that

$$I(C_1 C_2 \cdots C_{n-1} : C_0) = 0. \quad (24)$$

The data-processing inequality implies that

$$I(C_1 C_2 \cdots C_{n-1} : B' C_0) \geq I(C_1 C_2 \cdots C_{n-1} : B'). \quad (25)$$

From Eqs. (23)-(25) we obtain that

$$\begin{aligned} I(C_0 C_1 \cdots C_{n-1} : B') \\ \geq I(C_0 : B') + I(C_1 C_2 \cdots C_{n-1} : B'). \end{aligned}$$

After iterating these steps $n - 1$ times, we have

$$I(C : B') \geq \sum_{k=0}^{n-1} I(C_k : B'). \quad (26)$$

Since the system B_k is output by Bob after local operations on his system B' , applying the data-processing inequality, we obtain $I(C_k : B') \geq I(C_k : B_k)$, which from Eq. (26) implies Eq. (22).

Upper bound on P from quantum information causality

We show an upper bound on the success probability P in the QIC game from quantum information causality:

$$P \leq P', \quad (27)$$

where we define P' to be the maximum solution of the equation

$$h(P') + (1 - P') \log_2 3 = 2 \left(1 - \frac{m}{n}\right), \quad (28)$$

and $h(x) = -x \log_2 x - (1 - x) \log_2 (1 - x)$ denotes the binary entropy.

We notice that since Charlie's and Bob's systems are initially uncorrelated, the quantum information causality

bound (Eq. (4) of the main text) reduces to $I(C : B') \leq 2m$. Thus, from the bound given by Eq. (22) we have that

$$\sum_{k=0}^{n-1} I(C_k : B_k) \leq 2m. \quad (29)$$

Charlie initially prepares the qubits C_k and A_k in the singlet state $|\Psi^-\rangle_{C_k A_k}$, which after Alice's and Bob's operations is transformed into some state ω_k , now in the joint system $C_k B_k$. We have shown that in general we can consider ω_k to be of the form given by Eq. (7) of the main text:

$$\omega_k = \lambda_k \Psi^- + \frac{1 - \lambda_k}{3} (\Psi^+ + \Phi^+ + \Phi^-).$$

Thus, we have that $I(C_k : B_k) = 2 - S(\omega_k)$. Hence, from Eq. (29) we have that

$$\frac{1}{n} \sum_{k=0}^{n-1} S(\omega_k) \geq 2 \left(1 - \frac{m}{n}\right). \quad (30)$$

We define the state $\omega \equiv \sum_{k=0}^{n-1} \omega_k / n$. From the concavity of the von Neumann entropy [1], we obtain $S(\omega) \geq \sum_{k=0}^{n-1} S(\omega_k) / n$, which together with Eq. (30) implies

$$S(\omega) \geq 2 \left(1 - \frac{m}{n}\right). \quad (31)$$

From the definitions of P (Eq. (2) of the main text) and ω , and the form of ω_k (Eq. (7) of the main text) we have that

$$\omega = P \Psi^- + \frac{1 - P}{3} (\Psi^+ + \Phi^+ + \Phi^-), \quad (32)$$

which has von Neumann entropy $S(\omega) = h(P) + (1 - P) \log_2 3$, where $h(x) = -x \log_2 x - (1 - x) \log_2 (1 - x)$ is the binary entropy. Thus, from Eq. (31) we have that

$$h(P) + (1 - P) \log_2 3 \geq 2 \left(1 - \frac{m}{n}\right), \quad (33)$$

which implies Eq. (27). This can be seen as follows. The function $h(P) + (1 - P) \log_2 3$ corresponds to the Shannon entropy of a random variable taking four values, one with probability P and the others with probability $(1 - P)/3$ [1]. It increases monotonically in the range $[0, 1/4]$ and decreases monotonically in the range $[1/4, 1]$. It takes the values $\log_2 3$ at $P = 0$ and $P = 0.609$, 2 at $P = 1/4$ and 0 at $P = 1$. If $2(1 - m/n) \geq \log_2 3$, Eq. (28) has two solutions, one in the range $[0, 1/4]$ and the other one in the range $[1/4, 0.609]$. Otherwise, Eq. (28) has a single solution in the range $(0.609, 1]$. Therefore, the maximum solution of Eq. (28) is in the range $[1/4, 1]$. Since in this range the function $h(P) + (1 - P) \log_2 3$ is monotonically decreasing, Eq. (33) implies Eq. (27).

In particular, we can easily see from Eq. (31) that if $m < n$ then $S(\omega) > 0$. Therefore, in this case ω cannot be a perfect singlet, which from Eq. (32) implies that $P < 1$.

Upper bound on Q for nonlocal strategies

We have obtained an upper bound on the success probability Q in the IC-2 game, defined in the main text, for a particular class of strategies in the case $m = 1$:

$$Q \leq Q', \quad (34)$$

where $Q' \equiv (1 + 3n^{-1/2})/4$. The considered class of strategies is the following.

Nonlocal strategies in the IC-2 game. Alice and Bob share an entangled state $|\psi\rangle \in \mathcal{H}$. They perform a local projective measurement on their part of $|\psi\rangle$. Alice chooses her measurement according to her value of $x \equiv (x_0, x_1, \dots, x_{n-1})$. Recall that $x_j \equiv (x_j^0, x_j^1)$, for $j = 0, 1, \dots, n-1$. Bob chooses his measurement according to his number k . Their measurement outcomes are the two bit numbers (a_k^0, a_k^1) and (b_k^0, b_k^1) , respectively. Alice sends Bob her outcome. Bob outputs the two bit value $y_k \equiv (y_k^0, y_k^1)$, where $y_k^j = a_k^j \oplus b_k^j$, for $j = 0, 1$, and \oplus denotes sum modulo 2. The success probability is

$$Q = \frac{1}{n} \sum_{k=0}^{n-1} P(y_k^0 = x_k^0, y_k^1 = x_k^1).$$

This class of strategies is not general. For example, a more general strategy would be one in which Bob uses Alice's message in order to choose his measurement.

It can easily be computed that for $m = 1$ and $n \geq 50$, $P' < Q'$, where P' is defined by Eq. (6) of the main text. Therefore, the bound given by Eq. (34) cannot be achieved for $n \geq 50$, otherwise Eq. (5) of the main text, and hence quantum information causality, could be violated by a teleportation strategy achieving $P = Q'$.

Now we present the proof of Eq. (34). This is an extension of the one given in Ref. [27] for the IC-1 game. Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Alice and Bob measure their respective systems, A and B , in the orthonormal bases $\{|\nu_{r,s}^x\rangle\}_{r,s=0}^1$ and $\{|w_{t,u}^k\rangle\}_{t,u=0}^1$. After the measurement is completed, the state $|\psi\rangle$ projects into the state $|\nu_{a_k^0, a_k^1}^x\rangle |w_{b_k^0, b_k^1}^k\rangle$. We define the Hermitian operators

$$\begin{aligned} \hat{A}_x &\equiv \sum_{r=0}^1 \sum_{s=0}^1 (-1)^{r+s} |\nu_{r,s}^x\rangle \langle \nu_{r,s}^x|, \\ \hat{B}_k &\equiv \sum_{t=0}^1 \sum_{u=0}^1 (-1)^{t+u} |w_{t,u}^k\rangle \langle w_{t,u}^k|, \end{aligned}$$

acting on \mathcal{H}_A and \mathcal{H}_B , respectively. We also define $E_{x,k} \equiv (-1)^{x_k^0 + x_k^1} \langle \psi | \hat{A}_x \hat{B}_k | \psi \rangle$. Writing the state $|\psi\rangle$ in the basis $\{|\nu_{r,s}^x\rangle |w_{t,u}^k\rangle\}_{r,s,t,u=0}^1$, using that $y_k^j = a_k^j \oplus b_k^j$, for $j = 0, 1$, and noticing that x is a completely random

variable of 4^n possible values, it is easy to obtain that

$$\begin{aligned} \frac{1}{n} \sum_{k=0}^{n-1} [P(y_k^0 = x_k^0, y_k^1 = x_k^1) + P(y_k^0 \neq x_k^0, y_k^1 \neq x_k^1)] \\ = \frac{1}{2} \left(1 + \frac{1}{n4^n} \sum_{x,k} E_{x,k} \right). \end{aligned} \quad (35)$$

Following the procedure of Ref. [27], it is obtained that

$$\frac{1}{2} \left(1 + \frac{1}{n4^n} \sum_{x,k} E_{x,k} \right) \leq \frac{1}{2} \left(1 + \frac{1}{\sqrt{n}} \right),$$

which from Eq. (35) implies

$$\begin{aligned} \frac{1}{n} \sum_{k=0}^{n-1} [P(y_k^0 = x_k^0, y_k^1 = x_k^1) + P(y_k^0 \neq x_k^0, y_k^1 \neq x_k^1)] \\ \leq \frac{1}{2} \left(1 + \frac{1}{\sqrt{n}} \right). \end{aligned} \quad (36)$$

This bound is achievable due to a theorem by Tsirel'son [32], as shown in Ref. [27].

Following a similar procedure, by defining $E_{x,k}^j \equiv (-1)^{x_k^j} \langle \psi | \hat{A}_x^j \hat{B}_k^j | \psi \rangle$, for $j = 0, 1$, in terms of the operators

$$\begin{aligned} \hat{A}_x^0 &\equiv \sum_{r=0}^1 \sum_{s=0}^1 (-1)^r |\nu_{r,s}^x\rangle \langle \nu_{r,s}^x|, \\ \hat{B}_k^0 &\equiv \sum_{t=0}^1 \sum_{u=0}^1 (-1)^t |w_{t,u}^k\rangle \langle w_{t,u}^k|, \\ \hat{A}_x^1 &\equiv \sum_{r=0}^1 \sum_{s=0}^1 (-1)^s |\nu_{r,s}^x\rangle \langle \nu_{r,s}^x|, \\ \hat{B}_k^1 &\equiv \sum_{t=0}^1 \sum_{u=0}^1 (-1)^u |w_{t,u}^k\rangle \langle w_{t,u}^k|, \end{aligned}$$

it can be shown that

$$\begin{aligned} \frac{1}{n} \sum_{k=0}^{n-1} [P(y_k^0 = x_k^0, y_k^1 = x_k^1) + P(y_k^0 \neq x_k^0, y_k^1 \neq x_k^1)] \\ \leq \frac{1}{2} \left(1 + \frac{1}{\sqrt{n}} \right), \end{aligned} \quad (37)$$

and that

$$\begin{aligned} \frac{1}{n} \sum_{k=0}^{n-1} [P(y_k^0 = x_k^0, y_k^1 = x_k^1) + P(y_k^0 \neq x_k^0, y_k^1 = x_k^1)] \\ \leq \frac{1}{2} \left(1 + \frac{1}{\sqrt{n}} \right). \end{aligned} \quad (38)$$

Adding Eqs. (36) - (38), using normalization of probabilities and arranging terms we obtain that

$$\frac{1}{n} \sum_{k=0}^{n-1} P(y_k^0 = x_k^0, y_k^1 = x_k^1) \leq \frac{1}{4} \left(1 + \frac{3}{\sqrt{n}} \right),$$

as claimed.

Similarly to the procedure of Ref. [27], a theorem by Tsirel'son [32] can be used to show that there exist measurement bases defining the binary operators $\hat{A}_x, \hat{B}_k, \hat{A}_x^0, \hat{B}_k^0, \hat{A}_x^1, \hat{B}_k^1$ and bipartite states on which the measurements are performed that allow Alice and Bob to achieve the bounds given by Eqs. (36) - (38). However, this does not mean that the three bounds can be saturated simultaneously with the same state $|\psi\rangle$ and the measurements defined by the bases $\{|\nu_{r,s}^x\rangle\}_{r,s=0}^1$ and $\{|\omega_{t,u}^k\rangle\}_{t,u=0}^1$.

-
- [1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).
 - [2] W. K. Wootters and W. H. Zurek, *Nature* (London) **299**, 802 (1982).
 - [3] D. Dieks, *Phys. Lett. A* **92**, 271 (1982).
 - [4] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, *Phys. Rev. Lett.* **70**, 1895 (1993).
 - [5] C. H. Bennett and S. J. Wiesner, *Phys. Rev. Lett.* **69**, 2881 (1992).
 - [6] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India* (IEEE, New York, 1984) p. 175.
 - [7] A. K. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).
 - [8] J. Barrett, L. Hardy, and A. Kent, *Phys. Rev. Lett.* **95**, 010503 (2005).
 - [9] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).
 - [10] J. S. Bell, *Physics* **1**, 195 (1964).
 - [11] G. C. Ghirardi, A. Rimini, and T. Weber, *Lett. Nuov. Cim.* **27**, 293 (1980).
 - [12] A. S. Kholevo, *Probl. Inf. Transm.* **9**, 177 (1973).
 - [13] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, *Nature* (London) **461**, 1101 (2009).
 - [14] J. Allcock, N. Brunner, M. Pawłowski, and V. Scarani, *Phys. Rev. A* **80**, 040103 (2009).
 - [15] D. Cavalcanti, A. Salles, and V. Scarani, *Nat. Commun.* **1**, 136 (2010).
 - [16] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, *Phys. Rev. Lett.* **107**, 210403 (2011).
 - [17] T. H. Yang, D. Cavalcanti, M. L. Almeida, C. Teo, and V. Scarani, *New J. Phys.* **14**, 013061 (2012).
 - [18] B. S. Cirel'son, *Lett. Math. Phys.* **4**, 93 (1980).
 - [19] S. Popescu and D. Rohrlich, *Found. Phys.* **24**, 379 (1994).
 - [20] O. E. Lanford and D. W. Robinson, *J. Math. Phys.* **9**, 1120 (1968).
 - [21] H. Araki and E. H. Lieb, *Commun. Math. Phys.* **18**, 160 (1970).
 - [22] L. Hardy, arXiv:quant-ph/0101012.
 - [23] J. Barrett, *Phys. Rev. A* **75**, 032304 (2007).
 - [24] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, *Phys. Rev. Lett.* **99**, 240501 (2007).
 - [25] A. J. Short and S. Wehner, *New J. Phys.* **12**, 033023 (2010).
 - [26] H. Barnum, J. Barrett, L. O. Clark, M. Leifer, R. Spekkens, N. Stepanik, A. Wilce, and R. Wilke, *New J. Phys.* **12**, 033024 (2010).
 - [27] S. W. Al-Safi and A. J. Short, *Phys. Rev. A* **84**, 042323 (2011).
 - [28] L. Masanes, M. P. Mueller, R. Augusiak, and D. Perez-Garcia, arXiv:1208.0493.
 - [29] S. Wiesner, *SIGACT News* **15**, 78 (1983).
 - [30] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, *J. ACM* **49**, 496 (2002).
 - [31] M. Pawłowski and M. Żukowski, *Phys. Rev. A* **81**, 042326 (2010).
 - [32] B. S. Tsirel'son, *J. Sov. Math.* **36**, 557 (1987).